Abstract

Neurodevelopmental disorders (NDs) are types of condition that affect how the brain functions and alter neurological development, causing problems in societal, intellectual, also emotive functioning. Autism Spectrum Disorder (ASD) is a difficult neurodevelopmental condition categorized by a variety of signs and severity. ASD encompasses a range of behavioural and social abnormalities, leading to difficulties in societal skills, repetitive behaviours, speech, and nonverbal communication. Although at hand is no definitive cure for ASD, primary finding allows for precautionary measures. Early and precise finding is essential for timely intervention. This report explores the usage of deep learning methods to develop predictive models for ASD diagnosis using a multi-modal approach incorporating MRI and genome data. Various deep learning models, with Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), will be evaluated. A deep neural network (DNN) model for classification and feature selection. The classifier, pre-trained with the feature extraction methods, showed significant performance improvement. Pre-training the autoencoder with the DNN classifier further enhanced the discriminative representation. Early ASD diagnosis challenges addressed through gene expression classifiers. Gene expression classifier replicable across ages, races, and ethnicities. Classifier outperforms risk gene mutation classifier, potential for clinical translation. The results demonstrate that integrating these data sources significantly enhances predictive accuracy, offering a promising direction for future research and clinical applications. Recent research focuses on diagnosing ASD due to the absence of a specific biomarker. Traditional diagnostic methods, relying on behavioural observation and interviews, often lack accuracy. Moreover, neuroimaging shows minimal differences between ASD subjects and healthy controls, complicating its use for diagnosis. Consequently, machine learning-based approaches for diagnosing ASD are gaining popularity, utilizing features extracted from functional or structural MRI images. Considering the importance of age in structural image studies for ASD diagnosis, We can used the dataset, covering a wide age range. The proposed methods will contribute to building an early diagnosis model for ASD.

Table of Contents

Sr. No	Contents	Page No
	Acknowledgement	iii
	Abstract	iv
1	Title	1
2	Aims of Study	2
3	Study Background(Literature Review) and Significance of study	2
4	National Status	15
5	International Status	15
6	Objectives of Study	16
7	Study Hypothesis	17
8	Study Rationale	17
9	Research Methodology	18
10	Statistical plan	21
11	Expected Outcomes of Study	22
12	Limitation of Study	22
13	Bibliography	22

List of Figures

Figure No.	Title	Page No.
1	Proposed Methodology Architecture	21

List of Tables

Table No.	Title	Page No.
1	Activity Details	21

1. Title: Autism Spectrum Disorder Detection and Classification in Pediatric Neurodevelopment to Identify the Biomarkers

1.1 Introduction:

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition affecting individuals' perception and interaction, leading to challenges in social interactions and communication, repetitive behaviors, and unique learning and attention styles. Autism is more common in men besides prematurely born children, and it is mostly related to the genetic condition fragile X [1]. As per the The Centre for Disease Control and Prevention (CDC), ASD, a neurodevelopmental disorder characterized by brain development differences [2]. Autism prevalence in India is increasing, with one in every 68 youngsters diagnosed, with males being more affected than females the male to female ratio is about three to one., according to a 2021 Indian Journal of Pediatrics study. In a similar vein, the Centers for Disease Control and Prevention (CDC) estimate that 1 in 36 American children suffer from ASD. [2].

Accurate data on ASD frequency is challenging due to lack of knowledge, limited medical access, and societal stigma surrounding neurodevelopmental diseases. The increasing occurrence of ASD underscores the necessity for early detection also intervention, benefiting both patients and healthcare providers by enabling appropriate therapy prescription and reducing long-term costs.

With the increased use of machine learning-based models for illness prediction, it is now possible to identify early based on the variety of physiological and health parameters. Finding people with ASD is difficult because many other mental disorders have few symptoms that are very similar to those of people with ASD.

According to recent study, a complex mode of autism may be attributed to an increased risk due to many genetic variables. People who have certain genetic disorders, such as Tuberous Sclerosis and Fragile X Syndrome, are more likely to be identified as having autism. These situations, along with numerous uncommon genetic causes, collectively account for more than 30% of autism cases [5].

1.2 Challenges in ASD Diagnosis

- Diagnosing Autism Spectrum Disorder (ASD) requires input from multiple experts including psychiatrists, therapists, special educators, counsellors, and neurologists.
- Typically, this diagnostic process involves comprehensive assessment over 2–3 days, incorporating parent interviews and standardized questionnaires.

- Swift diagnosis is crucial to promptly initiate effective treatment for children with ASD, aiming for better outcomes.
- Delayed detection of ASD can result in irreversible consequences.
- Machine-learning methods are effective tools for analysing ASD severity accurately, facilitating informed decision-making in clinical settings.

2. Aims of Study:

This research study is aimed at various levels of findings and solutions. It mainly focused on Autism Spectrum Disorder (ASD) Detection and Classification in children to Identify different biomarkers where the data can be used as brain Imaging and genetic. Here are some potential aims:

- To identify specific neuroimaging biomarkers by analyzing structural and functional brain imaging data and genetic variants and gene expression profiles associated with ASD in Pediatric patients.
- To develop a robust diagnostic model that combines neuroimaging and genomic data to improve the accuracy of ASD detection and classification.
- To identify early biomarkers of ASD that can be detected in very young children or even infants, enabling earlier diagnosis and intervention strategies
- To evaluate the validity and reliability of identified biomarkers across diverse pediatric populations with ASD.
- Contributing to a comprehensive understanding of how brain imaging and genomic data can be used to identify biomarkers and improve clinical outcomes for children with ASD.

3. Study Background(Literature Review) and Significance of study

3.1 Literature Review:

The paper [1], gives information as early identification of autism helps young people receive therapy and improve social skills. Early assessment is aided by the Screening Tool for Autism in Toddlers (STAT). Content authority perception is impacted by an author's lack of credentials. The disadvantages and problems surrounding the early diagnosis of ASD are rarely discussed. Broad possibilities for therapy that don't take into account each patient's effectiveness or level of symptoms. Alternative viewpoints, such as the neurodiversity model, are lacking. A dataset called OpenCV is used to train and improve models. It takes advantage of the official assessment tool, the ADOS-G, to evaluate ASD. For ASD screening, the SCQ

consists of 40 yes-or-no questions. Improved outcomes and specific interventions are among the benefits of early diagnosis that arise from this. Labelling, stigmatization, over diagnosis and intricate diagnostic procedures are among the risks.

According to these different websites [2-4] that provides the statistics data related to ASD, A study published in the Indian Journal of Paediatrics in 2021 estimated that approximately 1 in 68 children in India are affected by autism. The condition is more common in boys than in girls, with a male-to-female ratio of about 3:1. As per Centers for Disease Control and Prevention (CDC) and National institute of Mental Health (NIH), the prevalence is given where data were collected from health and special education records of 8-year-old children living in 11 areas of the United States during 2020, 1 in 36 (2.8%) 8-year-old children were estimated to have ASD across the CDC surveillance sites in 2020.

In the website [5] given, Current research indicates that multiple genetic factors can contribute to an increased risk of autism in a complex manner. Individuals with specific genetic conditions like Fragile X Syndrome and Tuberous Sclerosis have a higher likelihood of being diagnosed with autism. These conditions, along with numerous rare genetic causes, collectively account for more than 30% of autism cases. As a result, genetic testing might be suggested following an autism diagnosis

This paper [6] study gives; use of Deep learning (DL) for tasks like classification and prediction. While numerous biomarkers have been identified through neuroimaging for diagnosing, classifying, predicting outcomes, and monitoring responses to treatment in neurodevelopmental disorders, they haven't been effectively integrated to yield clinically useful outcomes. DL offers a potential solution to these challenges, providing meaningful insights that could enhance clinical practices. This review examines the integration of neuroimaging and DL in the study of neurodevelopmental disorders, highlighting the advantages, challenges, and prospects of this research area. Although neuroimaging has identified several biomarkers, their analysis often fails to produce results that are significant in a clinical context. Given the high comorbidity rates of neurodevelopmental disorders, leveraging DL to analyse clinical data from existing diagnostic assessments could improve diagnostic accuracy and aid clinicians in making more informed decisions about diagnosis and treatment.

This paper [7] work utilized a deep neural network approach on a substantial dataset of brain images sourced from ABIDE to improve the diagnosis of Autism Spectrum Disorder (ASD), particularly in children. Our model integrates unsupervised learning with an autoencoders and supervised learning using convolutional neural networks. Through extensive validation and assessment, our algorithm surpasses individual classifiers. Experimental findings demonstrate that combining the autoencoder with convolutional neural networks yields superior performance, attaining an accuracy of 84.05% and an Area under the Curve (AUC) value of 0.78. Training the model with an autoencoder and CNN took about 55 minutes. The autoencoder was effective in feature selection, allowing the selected features to be used in various classification models. The autoencoder, when used with SVM, Random Forest, and KNN, achieved AUC scores of 0.69, 0.65, and 0.59, respectively. A major limitation of our study was the small dataset size, which consisted of 1,112 rs-fMRI images. Looking ahead, incorporating personal characteristics such as birth weight, maternal age, and family history might yield more promising results. Additional data like daily behavior surveys or short video clips of the participants could enhance autism detection by enriching the feature set. Advanced models like ResNet-50 could be applied to fMRI data using transfer learning techniques. Future research could also explore the effectiveness of other dimensionality reduction techniques like PCA, LDA, locally linear embedding, and Isomap in conjunction with various machine learning or deep learning classifiers. Improvement in KNN results might be achieved through Grid Search CV, which iterates over predefined hyperparameters to optimize model settings. Overall, we developed and assessed the effectiveness of an autoencoder in conjunction with other models, comparing these hybrid approaches to the performance of individual models to gauge the impact of enhanced feature representation.

In this study[8], They explore the application of various Machine Learning (ML) methods such as Naive Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF) for analyzing and predicting Autism Spectrum Disorder (ASD) in two distinct age groups: toddlers (ages 12 to 36 months) and adolescents (ages 12 to 18 years). Utilizing non-clinical ASD datasets, they conduct a comprehensive literature review and delve into the efficacy of these supervised ML algorithms. The focus is on evaluating the performance of these four prominent classifiers within the context of ASD screening, aiming to provide insights into their comparative strengths and limitations in predictive accuracy.

They evaluated three dataset splitting techniques and found that a 70-30 split between training and testing data yields the most accurate results. The comprehensive testing used several

metrics to gauge model performance on two ASD datasets, one focusing on toddlers and the other on adolescents. The Random Forest (RF) classifier emerged as the most effective, achieving the highest accuracies of 93.69% for toddlers and 93.33% for adolescents. Other metrics, including precision, recall, F1-score, and AUC, were also considered. In comparison, the SVM model recorded accuracies of 77.28% for the toddler group and 89.33% for adolescents. These findings underscore the potential of machine learning techniques in enhancing our understanding and diagnosis of autism across different age groups.

In their study [9], the authors collected clinically diagnostic and leukocyte RNA data from 240 toddlers, including both Autism Spectrum Disorder (ASD) and typically developing (TD) children. They split the dataset with 175 toddlers used for training and 65 for testing. The aim was to create diagnostic classifiers for ASD using gene expression data. They built 42,840 models with various combinations of 3,570 gene expression features and 12 classification methods. Among these, 742 models attained high AUC-ROC scores (≥ 0.8) on both the training and test sets.

Using weighted Bayesian model averaging, they created an ensemble classifier from these 742 models, which demonstrated robust performance with AUC-ROC scores ranging from 85% to 89% and AUC-PR scores from 84% to 92% for ASD diagnosis. Toddlers with collective scores above or below the mean showed related diagnostic and psychometric results, but those scoring below the mean had higher prenatal risk events compared to TD toddlers.

The genes identified by the ensemble model are implicated in cell sequence regulation, safe response, gene regulation, and signalling pathways such as PI3K-AKT, RAS, and Wnt. Moreover, targeted DNA sequencing of ASD risk genes in a subset of toddlers revealed comparable mutation rates between ASD and TD groups. However, classifiers based solely on these mutations performed at chance level (49%), whereas the leukocyte group gene expression classifier accurately diagnosed 88% of toddlers with mutations in ASD risk genes, regardless of ASD or TD status.

Overall, the ensemble gene expression classifier for ASD shows promising predictive ability across different demographics of toddlers, surpassing mutation-based classifiers and suggesting potential for clinical application.

This study [10] focusing on autism spectrum disorder (ASD), researchers explored genomics data alongside traditional clinical interviews for diagnostic purposes. Recognizing the superior performance of deep neural networks (DNNs) with complex and high-dimensional data, the

study applied deep learning techniques to genetic data collected from thousands of simplex families at risk for ASD. Following data preprocessing from the Simons Simplex Collection, the study identified significant common genetic variants using chi-square tests, which could either protect against or contribute to ASD. These variants were then used to develop a CNN based diagnostic classifier aimed at predicting autism.

Comparisons were made with shallow machine learning models and randomly selected common variants to assess performance. The study found that the identified contributory common variants were notably enriched on chromosome X, with chromosome Y also showing discriminatory potential in distinguishing autistic individuals from non-autistic individuals. Specific genes such as ARSD, MAGEB16, and MXRA5 were highlighted for their significant effects among these variants, leading to adaptations in screening algorithms to incorporate these findings.

The deep learning model achieved an impressive area under the receiver operating characteristic curve (AUC-ROC) of 0.955 and an accuracy of 88% in identifying autistic individuals compared to non-autistic individuals. This represented a substantial improvement of approximately 13% in classification accuracy over traditional autism screening tools.

In this paper [11], Data collected from real-world environments often includes irrelevant information, which can introduce noise and diminish model performance. Gene expression, carrying genetic information, is crucial for understanding disease associations, but irregular molecular interactions during transcription can distort gene expression patterns, complicating the identification of disease biomarkers. To tackle this challenge, a novel gene selection approach is proposed to pinpoint potential biomarkers from genomic data. This approach integrates Signal to Noise ratio alongside logistic sigmoid function, Hilbert-Schmidt Independence Criterion Lasso, and a regularized genetic algorithm for optimizing feature selection. The method is validated using a microarray gene expression dataset from the Gene Expression Omnibus repository, specifically focusing on autism spectrum disorder (ASD).

Through this approach, genes like FAM104B, CCNDBP1, H1F0, and ZER1 are identified as candidate biomarkers for ASD. The performance of the model is rigorously evaluated using various machine learning algorithms, demonstrating enhanced prediction accuracy for ASD at 97.62%, surpassing current methods. This classification holds potential as a valuable tool to aid medical experts in accurately diagnosing ASD, contributing to improved patient outcomes and treatment strategies.

In this research [12], researchers created a machine learning model to classify ASD using mRNA expression data from peripheral blood samples of 128 toddlers with ASD and 126 control subjects. By comparing the gene expression levels, we identified several differentially expressed genes (DEGs) such as UBE4B, SPATA2, and RBM3, which are primarily associated with immune-related pathways. Using LASSO regression, they selected 21 genes as key biomarkers, achieving an accuracy of 86%. Subsequently, a neural network model incorporating these 21 genes attained an impressive Area under the Curve (AUC) of 0.88. Our study indicates that these neurotransmitters and immune-related biomarkers have the probable to aid in the early finding of ASD. The mRNA expression profiles provide insights into the biological mechanisms underlying ASD in toddlers and highlight promising biomarkers for early detection. Nonetheless, further validation with larger sample sizes is necessary to confirm the reliability of these biomarkers.

Here the authors of the paper [13], introduce DeepASDPred, A novel ASD risk gene analysis using deep learning techniques begins with K-mer encoding to represent RNA sequences, integrated with gene expression values to create a feature matrix. Chi-square testing and logistic regression are combined to select optimal features for a binary classification model using CNNs and LSTM networks. Tenfold cross-validation shows this method outperforms current state-of-the-art techniques in identifying ASD risk RNA genes.

The experimental findings demonstrate DeepASDPred's superior performance in identifying ASD risk genes, using CNNs and LSTMs solely on RNA sequences and gene expression data without prior biological knowledge. Efficiency is improved by reducing data redundancy and speeding up training through feature selection with chi-square testing and logistic regression. Comparative analyses with three classifiers and existing methods confirm DeepASDPred's effectiveness.

The study is limited by the small sample size of gene expression data. Additionally, emerging research highlights the potential role of non-coding RNAs in ASD, suggesting the need to expand gene databases and investigate non-coding RNAs for predicting ASD risk genes. These efforts could significantly advance our understanding and prediction capabilities in ASD research.

In this paper [14] Authors discussed that, Neuroimaging genetics plays a crucial role in identifying genes associated with ASD that impact both the structure and function of the brain, thereby validating biological alterations through elucidating the genetic pathways and

mechanisms involved. Through the combination of artificial intelligence (AI) models with neuroimaging data, a framework can be established to enhance the accuracy of diagnoses and contribute to the recognition of potential initial indicative biomarkers for ASD. This analysis emphasizes the importance of neuroimaging genetics in advancing our comprehension of disrupted neurochemical systems and molecular pathways in ASD, as well as in the detection of structural, functional, and metabolic variations, thus laying the groundwork for the identification of novel biomarkers for early diagnosis. In autism diagnosis, AI leverages behavioural data from patients, analysing it against other datasets to provide an initial diagnosis and recommend suitable therapeutic options. The heterogeneity of ASDs means that simultaneous functional and anatomical changes can complicate accurate diagnosis. Current machine learning and deep learning methods for ASD classification using 3D and 4D spatio-temporal data are still emerging and need improvement. Combining structural, functional, and metabolic imaging data with 3D or 4D convolutional neural networks could create more accurate tools for early ASD diagnosis.

The research explores [15] autism spectrum disorder (ASD) and its associated conditions such as depression, anxiety, and epilepsy, which can affect various aspects of life like education, social interaction, and employment. The paper delves into utilizing machine learning methods for detecting Autism Spectrum Disorder (ASD) in individuals of varying ages, from toddlers to adults. It initiates with visualizing data to identify patterns, followed by a thorough examination involving thirteen machine learning models across two distinct datasets. Model performance is assessed using various statistical measures. Introducing two novel deep learning architectures, namely the CNN–LSTM–PSO and GRU–CNN models, it demonstrates enhanced accuracy through optimization techniques. Additionally, the study incorporates AI fairness strategies to address bias, encompassing outlier management, transformations, and SMOTE. Comparative analysis underscores improvements over previous research, while acknowledging limitations concerning data availability, model complexity, and training duration. Despite benefiting from adequate data points, future studies could benefit from more extensive datasets to refine their findings.

These authors' of this paper [16] gives, computational approaches using gene expression data have emerged as effective tools for early ASD prediction. Feature selection techniques significantly impact the performance of ASD prediction methods. Given the proliferation of high-dimensional ASD gene expression data and advancements in computational methods,

there is a need to evaluate various techniques in predicting ASD. This paper reviews and compares 22 feature selection methods for ASD prediction from gene expression data, classified into traditional (14 methods) and network-based (8 methods). Experimental findings indicate that network-based methods generally outperform traditional ones across all accuracy measures, including AUC, F1-score, and Matthews Correlation Coefficient.

This study [17] acknowledges that autism is influenced by mutations across multiple genes, which are dynamically expressed in gene expression microarray data. The research utilizes microarray data. There are 146 samples in the dataset, divided between individuals with autism and a control group, with each sample featuring 54,613 gene expression features. In order to identify 100 relevant genes apiece, the study uses four feature selection techniques, with the goal of reducing dimensionality without sacrificing generality, given the dimensionality difficulties of these features and data. After careful hyperparameter setup, these chosen features are included to a neural network classifier model for autism classification. A neural network model is used in conjunction with each feature selection technique to examine the influence of well-known genes on enhancing autism categorization. The neural network model is evaluated on untrained data after being trained over 10 times using accuracy and loss as evaluation metrics. The optimized framework strikes a balance between accuracy and loss, achieving an accuracy of 78.6% and a low loss of 0.694 for untrained data.

This paper [18], provides the details as Genetics plays a pivotal role in understanding Autism Spectrum Disorder (ASD), with potentially thousands of genes implicated in its risk when they undergo functional disturbances. Despite this, only about 100 genes are currently recognized as definitive "autism risk genes," leaving a significant gap in knowledge. To address this, extensive genetic studies are on-going, but they are costly and slow-moving. In this context, identifying potential ASD risk genes using existing data becomes crucial. This study reframes the discovery of autism risk genes as a machine learning challenge, leveraging genome-scale data to find novel genes with characteristics akin to recognised autism threat genes. The forecASD ensemble method integrates brain gene expression data, diverse network data, and prior autism predictors into a unified classifier, scoring each gene based on its role in autism. It has shown to be more effective than previous methods in trio-based sequencing studies. The genes highlighted by forecASD demonstrate its strength in identifying autism-related genes, with potential uses in gene discovery, pathway analysis, and differential expression studies.

In this research [19], a machine learning classification model was developed using mRNA expression data obtained from the peripheral blood of 128 toddlers diagnosed with ASD and 126 controls. Through analysis, differentially expressed genes (DEGs) were identified, including notable ones such as UBE4B, SPATA2, and RBM3, primarily associated with immune-related pathways. Using LASSO regression, 21 genes were selected as key biomarkers, achieving an accuracy of 86%. Subsequently, a neural network model based on these 21 genes attained an AUC of 0.88. The study indicates that these identified neurotransmitters and immune-related biomarkers hold promise for early ASD diagnosis. By shedding light on the mRNA expression profile, this research provides insights into the biological mechanisms of ASD in toddlers and proposes potential biomarkers for early detection. However, further validation with larger sample sizes is necessary to confirm the efficacy of these biomarkers.

This study[20] proposes as, Functional magnetic resonance imaging (fMRI) serves as a valuable non-invasive tool for measuring brain activity with high resolution, surpassing the capabilities of electroencephalography (EEG). Clear visualization of brain activity is crucial for identifying autism-related anomalies. This study focuses on predicting Autism Spectrum Disorder (ASD) at an early stage by integrating multiple behavioral activities and developmental measures using machine learning algorithms. Various prediction methods are explored, with a particular emphasis on neuroimaging and ultra-high-risk factors. Moderate accuracy in predicting ASD is achieved using developmental measures obtained at 14 months from multiple time points. The study proposes augmenting early prediction using the Mullen scale and employs a computational approach to fMRI analysis with an adaptive functioning classifier for machine learning algorithms. This approach yields improved classification in machine languages, demonstrated by enhanced accuracy with conservative methods.

The study [21] concentrates on employing machine learning to detect autism spectrum disorder (ASD). Early identification of ASD is crucial for preventing potential long-lasting impacts. The paper employs federated learning (FL) as a technique for precise ASD diagnosis. Local training of two distinct machine learning classifiers, logistic regression, and support vector machine, is conducted. The outcomes from these classifiers are sent to a central server for training a Meta classifier. The proposed model attains a 98% accuracy rate in detecting ASD among children and an 81% accuracy rate among adults

The paper [22] introduces a machine learning framework aimed at detecting Autism Spectrum Disorder (ASD) at an early stage. It employs various machine learning techniques and feature scaling methods, conducting experiments on four different ASD datasets covering different age groups. Through these experiments, it identifies the most effective classification methods and feature scaling approaches. Notably, Ada Boost achieves the highest accuracy rates of 99.25% for Toddlers and 97.95% for Children, while Linear Discriminant Analysis performs best with accuracy rates of 97.12% for Adolescents and 99.03% for Adults. Additionally, the paper conducts feature importance analysis using four selection techniques. Overall, the proposed framework demonstrates promising results in early ASD detection compared to existing methods.

In a study[23] utilizing machine learning techniques, Logistic Regression achieved the highest accuracy of 93.15% among five models tested: Logistic Regression (LR), Naïve Bayes (NB), Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Random Forest Classifier (RFC). The models were trained and tested using forward feature selection. The Q-CHAT-10 screening method for toddlers is particularly noteworthy in early detection. SVM followed with an accuracy of 93.84%, NB with 94.79%, KNN with 90.52%, and RFC with 81.52%. These findings underscore the potential of machine learning, especially Logistic Regression, in advancing ASD diagnosis by reducing time and enhancing accuracy in clinical settings.

This research[24] explores a novel approach to diagnose Autism Spectrum Disorder (ASD) in children using a combination of internal neurophysiological (EEG) and external behavioral (eye-tracking, ET) data. The proposed framework integrates these modalities through a stacked denoising autoencoder (SDAE) model in a two-step process. Initially, separate SDAEs are employed to learn features from EEG and ET data. Subsequently, a third SDAE fuses these learned features to enhance diagnostic accuracy by capturing correlations and complementarity between the modalities. The multimodal identification model aims to generate informative and discriminative feature representations, leveraging the strengths of both neurophysiological and behavioral data. Evaluations conducted on a dataset comprising 40 ASD children and 50 typically developing (TD) children demonstrate that the proposed method outperforms single-modality approaches and a basic feature-level fusion method. These findings suggest the potential of the framework to offer clinicians objective and precise diagnostic support for ASD.

This paper [25] introduced, A machine learning framework was employed to explore biomarkers associated with Autism Spectrum Disorder (ASD). Functional network topology derived from fMRI data served as the primary biomarker. Functional connectivity (FC) and various complex network metrics were utilized to aid in diagnosis. By combining these features, the framework achieved the highest accuracy in diagnosing ASD. Specific brain regions exhibited notable differences in ASD patients, highlighting their potential relevance in understanding ASD pathology. The network topology approach provided distinct insights into ASD pathology not captured by traditional methods. Future research should consider integrating multiple fMRI analysis approaches to further enhance predictive capabilities.

This study [26] investigates both adults and children diagnosed with Autism Spectrum Disorder (ASD). Their method involves three main stages: dataset preparation, analysis, and unsupervised classification. Results from the classification show a sensitivity of 78.6% and specificity of 82.47% for adults, and a sensitivity of 87.5% and specificity of 95.7% for children. The complexity of ASD diagnosis is compounded by varying severity levels and a wide spectrum of symptoms. From an initial set of 73 criteria, only 15 variables were deemed relevant for adults and 19 for children. The study exclusively focuses on analyzing ASD data in adults and children, excluding other age groups. It's noted that the dataset used in the study is not publicly accessible.

In this research [27], a recommender model is created using multiple classifiers to improve the accuracy of predicting Autism Spectrum Disorder (ASD). The study evaluates different machine learning algorithms to determine their effectiveness. It demonstrates that Decision Trees and Random Forests show enhanced performance compared to other algorithms when assessed using metrics such as accuracy, precision, recall, and F1-score.

In order to improve diagnosis, this research [28] seeks to determine the most important features and automate the diagnostic procedure using existing classification approaches. The study analysed ASD datasets across different age groups: toddlers, children, adolescents, and adults. To find the best classifier and feature set for each dataset, cutting edge feature selection and classification techniques were assessed. Experiments show that the multilayer perceptron (MLP) classifier achieves 100% accuracy with minimum amount of attributes across all age groups, consistently outperforming other benchmark classification methods. The 'relief F'

feature selection technique was identified as the most effective method for ranking the most significant attributes in all four ASD datasets.

The study [29] focused on identifying biomarkers of autism in young children using magneto encephalography (MEG). Researchers recorded brain activity in 30 children with autism and 30 typically developing peers while they watched cartoons. They examined neural oscillations using amplitude (power spectral density, PSD) and phase (preferred phase angle, PPA). A machine learning classifier performed better with PPA features (88% accuracy) than with PSD features (82%). A new fusion approach combining both PSD and PPA features reached average accuracies of 94% for feature-level fusion and 98% for score-level fusion. These findings provide new insights into autism's neural characteristics in young children, potentially enhancing diagnostic understanding and methods.

This study [30] investigates the feasibility of automated ASD screening in toddlers under 24 months old, focusing on specific behaviours observed through gaze analysis using eye-tracking devices. Using a matched pair’s experimental design with a series of test videos, the research evaluates various machine learning algorithms. These algorithms, totalling nine different learning schemes, were selected based on their relevance in previous related studies and their popularity in classification tasks. The findings indicate that several of the tested algorithms demonstrated significant performance in accurately identifying ASD-related behaviours from gaze patterns in toddlers.

This review paper [31] discusses ensemble methods, which combine multiple base models to create a stronger overall model that often outperforms individual models. Success in ensemble methods depends on how these base models are trained and how their predictions are aggregated. While deep learning models have significantly improved predictive accuracy across many domains by automatically extracting features from data, they also require extensive expertise to tune hyperparameters effectively, which can be a complex and time-consuming task. Recent research has focused on integrating ensemble learning with deep learning to address these challenges. However, most existing approaches primarily utilize simple ensemble methods, which have limitations. The review paper aims to provide a comprehensive overview of ensemble learning strategies, particularly in the context of deep learning. It explores various factors influencing the effectiveness of ensemble methods and categorizes research efforts that have applied ensemble learning across diverse domains.

This review [32] synthesizes how technology is utilized in diagnosing and treating Neurodevelopmental Disorders (NDD), focusing on ASD and ADHD. Conducted under Prospero registration CRD42018091156, the review screened 808 full-text articles from 7982 initial identifications, ultimately including 47 papers. It evaluates the clinical effectiveness, economic impact, and user acceptability of various technologies such as mobile apps, robots, gaming, and virtual reality in NDD assessment and treatment. While many studies demonstrate some effectiveness, there is a call for higher-quality research, including randomized controlled trials with longer follow-up and economic evaluations, as well as qualitative studies to assess user impact and feasibility.

This paper [33] reviews 43 studies that employ unsupervised machine learning techniques in Autism Spectrum Disorder (ASD) research. These techniques include k-means clustering, hierarchical clustering, model-based clustering, and self-organizing maps. The objective of the review is to offer a comprehensive survey of how unsupervised machine learning is currently utilized in ASD research. It aims to provide insight into the specific research questions addressed by these methods and their applications in understanding ASD-related data patterns and characteristics.

3.2 Research Gap

From the above Literature review, we will able to identify the gaps, which is listed below,

- A notable thing in ASD prediction research with machine learning is the lack of integration of multiple data sources, while existing studies often focus on single modalities like behavioral features, eye tracking markers, or neuroimaging data.
- Acquiring real-time data and addressing challenges associated with limited data availability are crucial tasks.
- Exploring ML techniques like transfer learning and deep learning can enhance ASD screening, early detection, severity assessment and treatment recommendation across different ages, highlighting the necessity across diverse demographic and clinical populations.
- Developing innovative techniques to integrate diverse data formats, such as behavioral observations, genetic data, brain imaging, EEG, and eye-tracking data, holds promise for a comprehension of ASD causes and enhancing predictive modeling accuracy.
- Exploring the use of genetic data for diagnosing ASD, this is important exactly to find the Risk genes.

- Integration of Multimodal Data: There is growing interest in combining brain imaging, genomic, and behavioral data to create more accurate and individualized diagnostic tools. Machine learning and AI are increasingly being used to analyze these complex datasets.
- Personalized Medicine: The ultimate goal is to develop personalized approaches to ASD diagnosis and treatment, tailored to the unique genetic and neurobiological profile of each individual.
- Severity level of the patient to identify and recommending the personalize treatment.

There are differences in the state of research in India and throughout the world regarding the identification and categorization of Autism Spectrum Disorder (ASD), particularly when it comes to the use of brain imaging and genetic data.

4. National Status:

Utilizing innovative technology like brain imaging and genetics, research on ASD is still in its early phases in India. While early autism detection and intervention are becoming increasingly important, the merging of neuroimaging and genetics for biomarker identification is still relatively new in Indian institutions.

ASD research is being conducted at prestigious institutions like CCMB, AIIMS, and NIMHANS, focusing on genetic and neuroimaging studies.

- Initiatives: To raise awareness, discover ASD early, and provide assistance, the Indian government has launched a number of projects. For instance, the Ministry of Social Justice and Empowerment's National Trust offers a range of services to people with developmental disabilities, including ASD.
- Obstacles: The primary obstacles are insufficient resources, restricted availability of sophisticated diagnostic instruments, and the requirement for enhanced specialized education for medical practitioners concerning neurodevelopmental diseases.

5. International Status:

The discipline of diagnosing and classifying autism spectrum disorder (ASD) in children through brain imaging and genomic data is fast developing on a global scale. An outline of the state of research and practice now is provided below:

- Research vs. Clinical Practice: Although genetic and brain imaging data are useful in research settings to comprehend the biology underlying ASD, they are not yet accepted as conventional methods of clinical diagnosis. Now, behavioral tests like the Autism

Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview-Revised (ADI-R) are the main tools used to diagnose ASD.

- Global Efforts: To enhance ASD identification and classification, a number of extensive worldwide consortia are collaborating to integrate genetic and brain imaging data. Examples of important participants in this endeavor are the Simons Foundation Autism Research Initiative (SFARI) and EU-AIMS (European Autism Interventions - A Multicentre Study for Developing New Medications).

6. Objectives of Study:

The objectives of the proposed work are to work on some of the issues reported by the researchers and research gap found from the literature survey:

- **To design the ensemble model for brain imaging to classify the ASD with severity grading.**

  Ensemble models are designed to efficiently compress and encode high-dimensional brain imaging data, such as MRI and fMRI, for accurate diagnosis of Autism Spectrum Disorder. Deep learning models using feature engineering and PCA can accurately diagnose ASD by analyzing brain-imaging data, utilizing deep learning's ability to interpret vast data.

- **To propose the model for identifying the optimal and discriminative gene expression biomarkers as cause of ASD.**

  The goal is to leverage the vast and detailed information contained within genetic data to identify potential genetic markers or patterns that are associated with ASD. This process can enhance the understanding of the genetic underpinnings of the disorder and improve diagnostic precision.

- **To develop the multimodal system by integrating the brain imaging and genomics data.**

  A multimodal deep learning framework for ASD diagnosis combines brain imaging and genomic data, aiming for more accurate, efficient, and potentially earlier diagnosis. This approach captures a comprehensive biological profile of ASD, potentially uncovering novel correlations and interactions not apparent when analyzing these modalities separately.

7. Hypothesis:

A study on detecting and classifying ASD in children using brain imaging and genomic data requires a clear research question, formulating hypotheses, and planning a methodologically sound approach.

Research Question: Will the combination of brain imaging and genomic data can enhance the precision of early detection and classification of Autism Spectrum Disorder in children?

Primary Hypothesis (H1): Compared to conventional diagnostic techniques based on behavioral assessments, the integration of brain imaging and genomic data will greatly increase the accuracy of ASD detection in children.

Secondary Hypothesis (H2): Compared to neurotypical children, children with ASD will exhibit unique neuroimaging and genetic profiles, enabling a more accurate classification of ASD subgroups.

Null Hypothesis (H0): When compared to conventional techniques, the accuracy of ASD identification and classification will not be increased by the integration of brain imaging and genomic data.

Research study design will includes the participants with the sample size and criteria with the study groups.

Inclusion Criteria: Children having an official diagnosis of autism spectrum disorder (ASD) as determined by the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview-Revised (ADI-R) are eligible to apply. Controls ought to be neurotypical and free of psychiatric or developmental problems.

Exclusion Criteria: Children with serious medical illnesses or other neurodevelopmental disorders (such as ADHD or intellectual incapacity) are excluded.

8. Study Rationale:

Early detection, precise diagnosis, and successful treatment of ASD are severely hampered by the disorder's variability in terms of both underlying biology and symptoms. Conventional diagnostic techniques, which mainly rely on behavioural assessments, cannot adequately capture the variability of presentations of ASD and frequently cause delays in diagnosis, especially in younger children.

Current Diagnostic Approaches:

Behavioral assessments, such as ADOS and ADI-R, are the gold standard for diagnosing Autism Spectrum Disorder (ASD), but they can be subjective, time-consuming, and may not accurately diagnose milder symptoms or ASD in very young children.

Need for Objective Biomarkers:

ASD has a strong biological basis, influenced by genetic and neurobiological factors. Identifying biomarkers for early detection and classification is crucial for improved diagnostic accuracy. Neuroimaging studies have identified structural differences in ASD brains, but their integration into clinical practice is limited. Advances in genomic data reveal numerous genetic variants associated with ASD, but no definitive marker is yet identified.

Multimodal Approach:

The integration of brain imaging and genomic data can reveal complex interactions between genetic predispositions and neurodevelopmental outcomes, potentially identifying distinct ASD subtypes. Advances in machine learning and AI can analyse large datasets, identifying patterns and interactions, and developing predictive models to accurately distinguish children with ASD from neurotypical controls.

The suggested investigation is both essential and appropriate. It seeks to overcome the shortcomings of present diagnostic techniques and open the door for an earlier, more precise, and more customized diagnosis and treatment of ASD by combining brain imaging and genomic data. The results of the study may have a significant impact on the area of autism research as well as the millions of people and families impacted by ASD worldwide.

9. Research Methodology:

This method involves a number of important steps. Initially, the system is loaded with the raw data for ASD. The aforementioned data is essential for doing further studies. For instance, data cleansing, normalization, and standardization are carried out when the data is imported. Normalization minimizes fluctuations in expression data, ensuring gene expression values are properly scaled for meaningful comparisons. This step reduces data dimensionality and focuses on genes most relevant to the classification problem, using various gene selection strategies based on analysis requirements.

The approach to enhancing ASD detection through AI involves a systematic methodology encompassing several key phases, each meticulously designed to ensure robustness and accuracy in the system's performance.

Phase I: Preparatory Step

Data collection and Pre-processing:

The ASD detection dataset is acquired from various sources and pre-processed to improve its quality for machine learning models. Neuroimages reveal structural and functional

information, eliminating scanner effects and artifacts. Preprocessing includes data acquisition, normalization, artifact removal, temporal filtering, and masking to enhance accuracy and reduce execution time. This process ensures diversity and relevance to the task.

Preprocessing genetic data involves QC checks, variant calling algorithms, quality filtering, batch effects correction, feature selection, and integrative methods. These steps ensure high quality, bias-free genetic data, enhancing the accuracy and reproducibility of genetic studies.

Data Description:

ABIDE is a collaboration between 17 international imaging sites to share neuroimaging data from 539 individuals with Autism Spectrum Disorder (ASD) and 573 healthy controls. The dataset includes structural and resting state functional MRI data and detailed phenotypic information. Gene expression data related to ASD was sourced from the Gene Expression Omnibus (GEO), managed by the NCBI. The dataset includes 16,383 genes and 569 samples, providing insights into gene expression patterns in AD, MCI, and healthy cognition. The data was normalized and used for analysis.

Phase II: Feature Extraction/Selection Techniques:

Feature extraction transforms original features into a more informative and compact set. It aims to capture essential information in a lower-dimensional space and can be categorized into linear and nonlinear methods.

- Linear methods, like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), use linear transformations for feature extraction. PCA identifies principal components that capture the most modification, while LDA finds projections that boost class separability.
- Nonlinear methods, such as Kernel PCA and Autoencoders, use nonlinear transformations for feature extraction. Kernel PCA maps data to a higher-dimensional space with kernel functions to find principal components, while Autoencoders use neural networks to compress data into a lower-dimensional space and reconstruct it to the original form.

Feature selection is a process that chooses a subset of features from the original features so that the feature space is optimally reduced according to a certain criterion.

Feature selection algorithms generally fall into three categories: filter methods, wrapper methods, and embedded methods. The role of feature assortment is like,

1. Decrease the feature space dimensionality.
2. Accelerate the learning algorithm.
3. Enhance the analytical accuracy of classification.
4. Improve the clarity of learning results.

Phase III: Model Selection and Training

The next phase focuses on selecting appropriate Deep learning algorithms for ASD detection and training them using the pre-processed dataset. We choose to implement and train classifier. The dataset is split into training, validation, and testing sets to facilitate model training and evaluation. Hyperparameter tuning is performed using techniques such as grid search with cross-validation to optimize the performance of each algorithm.

Phase IV: Performance Evaluation

Repeated experiments and evaluation metrics are used to determine model performance in detecting ASD. These metrics provide insights into the effectiveness of each algorithm, and visualizations like confusion matrices and ROC curves are generated for further analysis. Evaluation metrics including accuracy, precision, recall, F1-score, Receiver Operating Characteristic (ROC) curve, and Area Under the Curve (AUC), Sensitivity, Specificity, MCC, etc.

Their definitions are listed as follows:

$$Accuracy = \frac{TN + TP}{TN + FP + TP + FN} \qquad Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN} \qquad F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

TP represents true positives, TN true negatives, FP false positives, and FN false negatives. ROC AUC and PR AUC were also used to evaluate model performance. The ROC curve plots the true positive rate against the false positive rate at various thresholds, while the PR curve shows precision versus sensitivity. Higher MCC, ROC AUC, and PR AUC values, closer to 1, indicate better model performance.

Phase V: Model Validation and Selection

The final phase validates algorithmic predictions for system robustness, analyzing performance metrics to identify the most effective algorithm for ASD detection. Cross-validation ensures objective evaluation, with the most accurate algorithm selected for real-world deployment, potentially influencing future research and clinical applications.

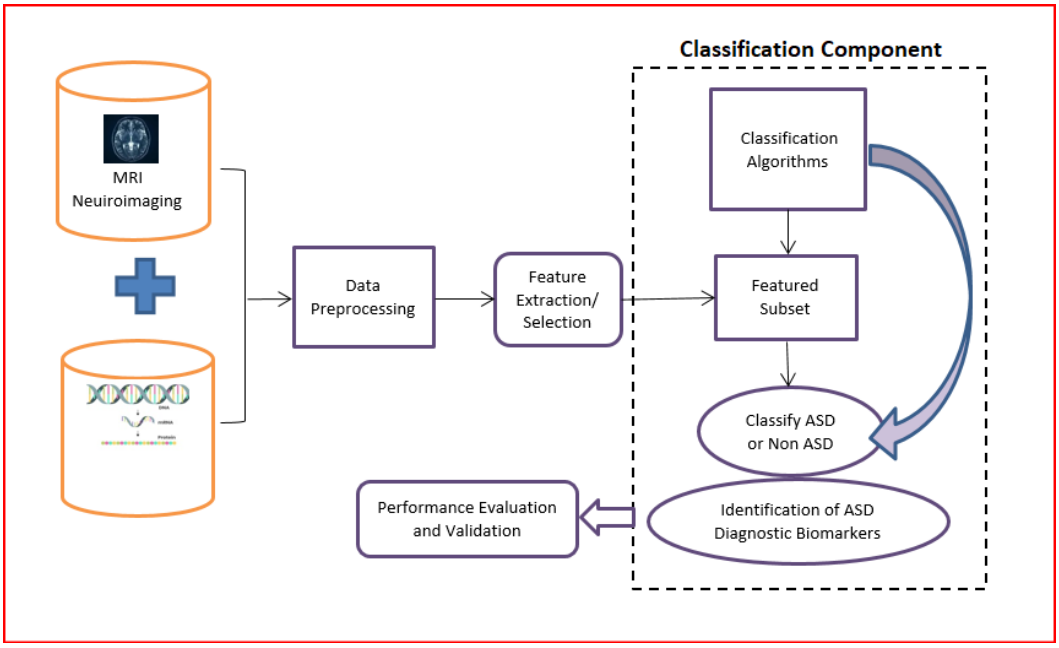

Fig 1.Praposed Methodology Architecture

The multimodal approach involves collecting brain imaging and genomic data, preprocessing, analyzing, and selecting relevant features. The combined data is then input into a model, trained, validated, and tested using cross-validation. Performance metrics like accuracy, precision, recall, and AUC are evaluated, highlighting key biomarkers and significant features for clinical implications.

10. Statistical plan:

The details of activity are shown in table 1.

Table 1: Activity details

Sr.No	Task/Stage	Plan Start	Plan Duration	Actual Start	Actual Duration
1	Literature Review & Research Design	Month 1	6 months	Month 1	6 months
2	Data Collection & Preprocessing data	Month 7	6 months	Month7	6 months
3	Data Analysis and Attribute Selection	Month 13	6 months	Month 13	6 months
4	Model Development	Month 19	6 months	Month 19	6 months
5	Result Validation and Evaluation	Month 25	6 months	Month 25	6 months
6	Thesis Writing	Month 31	6 months	Month 31	6 months

11. Expected Outcomes of Study:

- Enhancement of ASD detection through AI by evaluating different learning algorithms on curated dataset with meticulous preprocessing, which ensures data integrity and reliability, for subsequent analysis.
- The ensemble-learning framework consolidates brain imaging data, enabling comprehensive ASD analysis and prediction with severity levels.
- The propose model using the gene expression data to classify the ASD will give us the optimal and discriminative gene expression biomarkers/risk genes as cause of it.
- Integrating MRI and genomic data in a multimodal deep learning framework offers a promising approach to revolutionize ASD diagnosis, enhancing both accuracy and efficiency, helps identify subtle patterns and biomarkers essential for ASD diagnosis, offering insights into the disorder's mechanisms.

12. Limitation of Study:

The integration of genetic and brain imaging data for the diagnosis and categorization of autism spectrum disorder (ASD) holds great promise for research and therapeutic applications. Key areas where this research can advance and have an impact on the field are listed below:

- Severity assessment and treatment recommendation across different ages, highlighting the necessity across diverse demographic and clinical populations.
- Subsequent investigations may concentrate on customizing treatments according to a child's distinct neuroimaging and genetic information, resulting in more efficient and personalized therapy regimens.
- Multi-Omics Approaches: By combining genomic data with information from other omics fields, such as epigenomics (DNA methylation), transcriptomics (gene expression), proteomics (protein levels), and metabolomics (metabolite levels), a more thorough understanding of the biological pathways underlying ASD may be possible.

Bibliography

[1] Okoye C, Obialo-Ibeawuchi CM, Obajeun OA, Sarwar S, Tawfik C, Waleed MS, Wasim AU, Mohamoud I, Afolayan AY, Mbaezue RN. Early Diagnosis of Autism Spectrum Disorder: A Review and Analysis of the Risks and Benefits. Cureus. 2023 Aug 9;15(8):e43226. doi: 10.7759/cureus.43226. PMID: 37692637; PMCID: PMC10491411.

[2] Data & Statistics on Autism Spectrum Disorder, https://www.cdc.gov/autism/dataresearch/?CDC_AAref_Val=https://www.cdc.gov/ncbddd/autism/data.html

[3]Prevalence of ASD in India https://www.indiaautismcenter.org/early-detection-and-diagnosis-of-autism-in-india-importance-and-challenges/

[4] Prevalence of ASD, https://www.nimh.nih.gov/health/statistics/autism-spectrum-disorder-asd

[5] Autism Spectrum Disorder, https://www.psychiatry.org/patients-families/autism/what-is-autism-spectrum-disorder

[6] Song JW, Yoon NR, Jang SM, Lee GY, Kim BN. Neuroimaging-Based Deep Learning in Autism Spectrum Disorder and Attention-Deficit/Hyperactivity Disorder. Soa Chongsonyon Chongsin Uihak. 2020 Jul 1;31(3):97-104. doi: 10.5765/jkacap.200021. PMID: 32665754; PMCID: PMC7350542.

[7] Sewani H, Kashef R. An Autoencoder-Based Deep Learning Classifier for Efficient Diagnosis of Autism. Children (Basel). 2020 Oct 14;7(10):182. doi: 10.3390/children7100182. PMID: 33066454; PMCID: PMC7602176.

[8] Jyotismita Talukdar, Deba Kanta Gogoi, Thipendra P. Singh, A comparative assessment of most widely used machine learning classifiers for analysing and classifying autism spectrum disorder in toddlers and adolescents, 2772-4425/© 2023 The Author(s). Published by Elsevier Inc., https://doi.org/10.1016/j.health.2023.100178.

[9] Bao, B., Zahiri, J., Gazestani, V.H. *et al.* A predictive ensemble classifier for the gene expression diagnosis of ASD at ages 1 to 4 years. *Mol Psychiatry* **28**, 822–833 (2023). https://doi.org/10.1038/s41380-022-01826-x

[10] Wang H, Avillach P. Diagnostic Classification and Prognostic Prediction Using Common Genetic Variants in Autism Spectrum Disorder: Genotype-Based Deep Learning. JMIR Med Inform. 2021 Apr 7;9(4):e24754. doi: 10.2196/24754. PMID: 33714937; PMCID: PMC8060867.

[11] Sekaran, K., Sudha, M. Predicting autism spectrum disorder from associative genetic markers of phenotypic groups using machine learning. *J Ambient Intell Human Comput* **12**, 3257–3270 (2021). https://doi.org/10.1007/s12652-020-02155-z

[12] Tang H, Liang J, Chai K, Gu H, Ye W, Cao P, Chen S, Shen D. Artificial intelligence and bioinformatics analyze markers of children's transcriptional genome to predict autism spectrum disorder. Front Neurol. 2023 Jul 17;14:1203375. doi: 10.3389/fneur.2023.1203375. PMID: 37528852; PMCID: PMC10390071.

[13] Fan, Y., Xiong, H. & Sun, G. DeepASDPred: a CNN-LSTM-based deep learning method for Autism spectrum disorders risk RNA identification. *BMC Bioinformatics* **24**, 261 (2023). https://doi.org/10.1186/s12859-023-05378-x

[14] Nisar, S., Haris, M. Neuroimaging genetics approaches to identify new biomarkers for the early diagnosis of autism spectrum disorder. Mol Psychiatry 28, 4995–5008 (2023). https://doi.org/10.1038/s41380-023-02060-9

[15] Priyadarshini I. Autism Screening in Toddlers and Adults Using Deep Learning and Fair AI Techniques. *Future Internet*. 2023; 15(9):292. https://doi.org/10.3390/fi15090292

[16] Zhang, Junpeng, et al. "Computational Methods for Predicting Autism Spectrum Disorder from Gene Expression Data." International Conference on Advanced Data Mining and Applications. Cham: Springer International Publishing, 2020.

[17] Roopa, B. S., and R. Manjunatha Prasad. "A Selection of an Optimal Framework Identifying the Prominent Autism Risk Gene Biomarkers from Gene Expression Data Using Neural Network." SN Computer Science 2.4 (2021): 241.

[18] Brueggeman, L., Koomar, T. & Michaelson, J.J. Forecasting risk gene discovery in autism with machine learning and genome-scale data. Sci Rep 10, 4569 (2020). https://doi.org/10.1038/s41598-020-61288-5

[19] Tang H, Liang J, Chai K, Gu H, Ye W, Cao P, Chen S and Shen D (2023) Artificial intelligence and bioinformatics analyze markers of children's transcriptional genome to predict autism spectrum disorder. Front. Neurol. 14:1203375. doi: 10.3389/fneur.2023.1203375

[20] Karunakaran, P., and Yasir Babiker Hamdan. "Early prediction of autism spectrum disorder by computational approaches to fMRI analysis with early learning technique." Journal of Artificial Intelligence 2.04 (2020): 207-216.

[21] Muhammad Shoaib Farooq, Rabia Tehseen, Maidah Sabir & Zabihullah Atal "Detection of autism spectrum disorder (ASD) in children and adults using machine learning " Scientific Reports, (2023) 13:9605 https://doi.org/10.1038/s41598-023-35910-1

[22] S. M. Mahedy Hasan, M. P. Uddin, M. A. Mamun, M. I. Sharif, A. Ulhaq and G. Krishnamoorthy, "A Machine Learning Framework for Early-Stage Detection of Autism Spectrum Disorders," in *IEEE Access*, vol. 11, pp. 15038-15057, 2023, doi: 10.1109/ACCESS.2022.3232490.

[23] Vakadkar, K., Purkayastha, D. & Krishnan, D. Detection of Autism Spectrum Disorder in Children Using Machine Learning Techniques. *SN COMPUT. SCI.* **2**, 386 (2021). https://doi.org/10.1007/s42979-021-00776-5

[24] J. Han, G. Jiang, G. Ouyang and X. Li, "A Multimodal Approach for Identifying Autism Spectrum Disorders in Children," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 2003-2011, 2022, doi: 10.1109/TNSRE.2022.3192431.

[25] N. Chaitra , P.A. Vijaya , Gopikrishna Deshpande, "Diagnostic prediction of autism spectrum disorder using complex network measures in a machine learning framework", https://doi.org/10.1016/j.bspc.2020.102099, Biomedical Signal Processing and Control journal homepage: www.elsevier.com/locate/bspc

[26] Ammar I. Shihab, Faten A. Dawood , and Ali H. Kashmar, "Data Analysis and Classification of Autism Spectrum Disorder Using Principal Component Analysis" , Hindawi Advances in Bioinformatics Volume 2020, Article ID 3407907, 8 pages https://doi.org/10.1155/2020/3407907.

[27] Anita Vikram Shinde, Dipti Durgesh Patil, A Multi-Classifier-Based Recommender System for Early Autism Spectrum Disorder Detection using Machine Learning, Healthcare Analytics, Volume 4, 2023, 100211, ISSN 2772-4425, https://doi.org/10.1016/j.health.2023.100211.

[28] Hossain, M.D., Kabir, M.A., Anwar, A. *et al.* Detecting autism spectrum disorder using machine learning techniques. *Health Inf Sci Syst* **9**, 17 (2021). https://doi.org/10.1007/s13755-021-00145-9

[29] Barik, K., Watanabe, K., Bhattacharya, J. et al. A Fusion-Based Machine Learning Approach for Autism Detection in Young Children Using Magnetoencephalography Signals. J Autism Dev Disord 53, 4830–4848 (2023). https://doi.org/10.1007/s10803-022-05767-w

[30] Fernandez-Lanvin, D., Gonzalez-Rodriguez, M., De-Andres, J. *et al.* Towards an automatic early screening system for autism spectrum disorder in toddlers based on eye-tracking. *Multimed Tools Appl* (2023). https://doi.org/10.1007/s11042-023-17694-8

[31] Ammar Mohammed, Rania Kora, A comprehensive review on ensemble deep learning: Opportunities and challenges, Journal of King Saud University - Computer and Information Sciences, ScienceDirect. Volume 35, Issue 2, 2023, Pages 757-774, ISSN 1319-1578, https://doi.org/10.1016/j.jksuci.2023.01.014.

[32] Althea Z. Valentine, Beverley J. Brown, Madeleine J. Groom, Emma Young, Chris Hollis, Charlotte L. Hall, A systematic review evaluating the implementation of technologies to assess, monitor and treat neurodevelopmental disorders: A map of the current evidence, Clinical Psychology Review, Volume 80, 2020, 101870, ISSN 0272-7358, https://doi.org/10.1016/j.cpr.2020.101870.

[33] Parlett-Pelleriti, C.M., Stevens, E., Dixon, D. *et al*. Applications of Unsupervised Machine Learning in Autism Spectrum Disorder Research: a Review. *Rev J Autism Dev Disord* **10**, 406–421 (2023). https://doi.org/10.1007/s40489-021-00299-y

[34] Abdelwahab, Mahmoud M., Khamis A. Al-Karawi, and H. E. Semary. "Integrating gene selection and deep learning for enhanced Autisms' disease prediction: a comparative study using microarray data." AIMS Mathematics 9, no. 7 (2024).

[35] Y. Wang, J. Liu, Y. Xiang, J. Wang, Q. Chen, J. Chong, MAGE: automatic diagnosis of autism spectrum disorders using multi-atlas graph convolutional networks and ensemble learning, Neurocomputing (2021), doi: https://doi.org/10.1016/j.neucom.2020.06.152